Niveau(x) d'enseignement: 0 Numéro EDUQ : 8170638
Centre d'intérêt: 35

La Haye, Jacques.

ELMOPE '70-75 : rapport d'étape no 5 : montage du fichier "Mobilité du personnel" : opération du montage automatique, validité du processus / Jacques La Haye ; [pour le] Ministère de l'éducation, Démographie scolaire. -- [Québec] : Le Service, 1977.
32 f. : ill. ; 28 cm. -- (Document de travail / Service de la démographie scolaire du Québec ; 9.45)

Titre de la couv.

Cote Dewey: 371.1'009714/L183e Cote LC: LB/2833.4/.C38

Descripteur(s): *Mobilité professionnelle *Personnel administratif *Personnel enseignant *Personnel non enseignant *Banque de données

Identificateur(s) de nom: *ELMOPE (Projet) *Mobilité du personnel (Fichier)

Identificateur(s) de sujet: *Fichier informatisé

Le projet "ELMOPE 70-75" comporte le montage d'un fichier informatique supportant une information sur les membres du personnel scolaire qui soit à la fois individualisée et longitudinale. Le cinquième rapport d'étape décrit les opérations qui ont permis la constitution du fichier "Mobilité du personnel" à partir des six fichiers "Transit". Les trois chapitres présentent la logique des opérations de montage, la description du montage automatique de la banque de données et discutent de la validité du processus de montage.

L'originalité de la banque "Mobilité du personnel" consiste en ce que tout y est ramené en fonction de l'individu et ce pour les années 1970-1971 à 1975-1976. Ainsi, au lieu de traiter 604173 dossiers pour cette période, la banque en traite 154923 correspondant tous à des individus différents. Chaque dossier comporte une partie fixe établissant l'identité de l'individu et une série de "remarques" contenant les informations propres à chaque année. Les doubles sont donc éliminés, le phénomène des faux mouvements contrôlé et l'information standardisée sur une période de six ans.

Type de document: r
Notice(s) connexe(s): 8170634, 8170635, 8170636, 8170637
Auteur de l'analyse: ar

E D U Q 8170638

Gouvernement du Québec
Ministère de l'Education

Démographie scolaire
Document de travail D.T.9.45

"ELMOPE ' 70-75"

Montage du Fichier "Mobilité du Personnel"

- Opération du montage automatique
- Validité du processus

Rapport d'étape no 5

Juillet 1977 Jacques La Haye

E D U Q 8170638

TABLE DES MATIERES

page

INTRODUCTION ---------- 1

CHAPITRE I ---------- 3

Logique des opérations de montage

1. Triage des fichiers "Transit" par individu ---------- 3

2. Mise en correspondance des fichiers "Transit" ---------- 4

CHAPITRE II ---------- 8

Montage automatique de la Banque

1. Les éléments de montage ---------- 9

2. L'enchaînement des opérations ---------- 11

3. Résultats du montage ---------- 16

CHAPITRE III ---------- 23

Validité du processus de montage

1. Le code de résolution ---------- 23

2. Le code d'authenticité ---------- 25

CONCLUSION ---------- 29

ANNEXE ---------- 31

Liste des rapports d'étape couvrant le projet "ELMOPE '70-75" ---------- 32

LISTE DES TABLEAUX

page

TABLEAU 1 Schéma des opérations de montage de la Banque "Mobilité du Personnel", à partir des fichiers "Transit" annuels. Période 1970-71 à 1975-76. Montage sur le code No d'Assurance-Sociale-Sexe (N.A.S.S.), ETAPE 1. ------ 12

TABLEAU 2 Schéma des opérations de jumelage des dossiers ayant un N.A.S.S. défectueux (absent ou incomplet) sur la Banque "Mobilité du Personnel", ETAPE II. ---------- 14

TABLEAU 3 Schéma des opérations de montage de la Banque "Mobilité du Personnel", à partir des fichiers "Transit" annuels. Période 1970-71 à 1975-76. Montage sur le code No d'Assurance-Sociale-Sexe (N.A.S.S.), ETAPE I. Nombre de dossiers impliqués. ---------- 17

Tableau 4 Schéma des opérations de jumelage des dossiers ayant un N.A.S.S. défectueux (absent ou incomplet) sur la Banque "Mobilité du Personnel", ETAPE II. Nombre de dossiers impliqués. ---------- 18

Tableau 5 Etat de l'activité des individus de la Banque "Mobilité du Personnel" pour les six années scolaires de la période 1970-71 à 1975-76 (activité étant synonyme de "inscription au fichier"). ---------- 20

Tableau 6 Statisques sur l'utilisation des éléments de montage concernant les dossiers de la Banque "Mobilité du Personnel", au cours des opérations de montage automatique de la Banque. ----- 24

INTRODUCTION

Depuis la mise en branle de notre recherche sur le personnel scolaire, connue sous le nom de "projet ELMOPE", tous nos efforts ont été canalisés vers la préparation d'une banque de données qui nous ouvre la possibilité d'aborder la question de la mobilité du personnel à travers une approche longitudinale, c'est-à-dire d'une banque qui en arrive à nous permettre de suivre les membres du personnel scolaire, individu par individu, dans les divers réseaux.

Jusqu'à maintenant, quatre rapports d'étape[1] ont permis aux intéressés de suivre l'évolution du travail. On se rappelera que la création d'une telle banque d'information, qui devait assurer un lien de continuité entre les recensements annuels du personnel scolaire, était rendue complexe par le jeu de trois éléments principaux, à savoir: a) les modifications subies au cours des années par le formulaire individuel, b) l'absence de cohérence dans le traitement informatique des données et c) le sous-enregistrement des effectifs du personnel.

C'est ainsi que, avant de nous lancer dans le montage proprement dit de la banque, nous avons dû effectuer un important travail d'analyse des variables en présence, et ce juste-

---

(1) Voir la liste en annexe.

ment pour déterminer dans quelle mesure chacune d'elles pouvait être susceptible d'assurer la continuité d'une étude valable sur les six années de la période couverte (1970-71 à 1975-76) (voir les D.T. 9.36 et 9.39). Par la même occasion, nous avons pu dégager les faiblesses de chaque variable et, de là, définir le traitement approprié à chacune (voir D.T. 9.41). Ce sont finalement ces variables, sélectionnées et standardisées, qui furent placés en attentes sur les fichiers temporaires "Transit".

Dans ce cinquième rapport d'étape, nous voulons montrer précisément comment nous sommes alors passé de la série des six fichiers annuels "Transit" à un seul superfichier: la banque "Mobilité du Personnel", et ce, d'une façon purement automatique. Dans un premier temps, nous présenterons la logique des opérations de montage. Puis, nous montrerons comment les choses se sont déroulées en pratique, c'est-à-dire comment s'est opéré le montage automatique de la banque. Enfin, dans un dernier temps, nous discuterons de la validité d'un tel processus.

CHAPITRE I

LOGIQUE DES OPERATIONS DE MONTAGE

Le passage des fichiers annuels "Transit" à un seul fichier comporte deux phases d'opération:

- le triage des fichiers "Transit" par individu;
- la mise en correspondance des fichiers "Transit".

1. Triage des fichiers "Transit" par individu

On se rappellera que si un individu travaillait pour plus d'un organisme d'enseignement, il devait remplir un formulaire individuel pour chacun des organismes. C'est donc dire qu'un individu qui travaillait pour trois (3) organismes, par exemple, se retrouvait au fichier de base (Personnel-Maître) et, par conséquent, au fichier "Transit", avec trois enregistrements; en d'autres mots, qu'il était compté trois fois.

Le triage des fichiers "Transit" par individu vise donc précisément à éliminer ces pseudo-individus tout simplement en consignant l'information fournie par chacun des enregistrements dans un seul dossier. L'information de chacun des

enregistrements est alors conservée intacte, mais ne renvoie qu'à une seule personne.

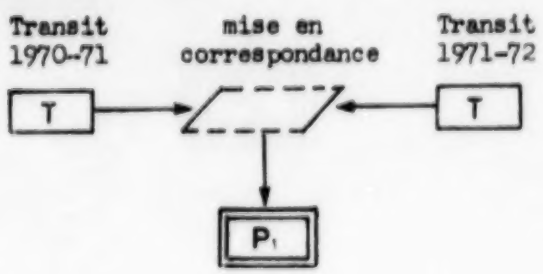
Cette opération, que nous appelons "jumelage intra-Transit", se répète pour chacun des six fichiers "Transit". Ce triage permet donc d'obtenir des "Transit" épurés où chaque individu n'est présent qu'une seule fois.

2. Mise en correspondance des fichiers "Transit"

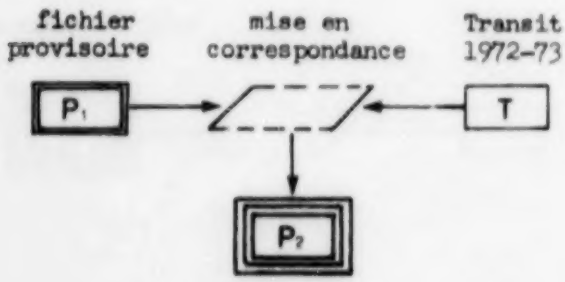
Des deux phases d'opération, la mise en correspondance constitue sans aucun doute la plus importante et aussi la plus délicate. Cette phase consiste essentiellement à fouiller dans chacun des fichiers "Transit" (épurés, bien entendu) pour aller chercher l'information se rapportant à un individu donné et à la consigner au dossier de cet individu.

L'opération comporte cinq (5) étapes successives:

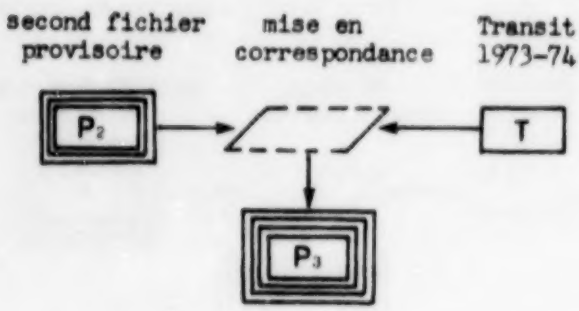
1- mise en correspondance des deux premiers fichiers "Transit" (1970-71 et 1971-72): de là, création d'un premier fichier provisoire (P_1);



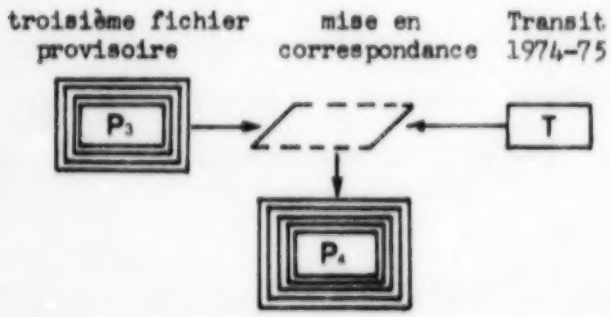
2- mise en correspondance du premier fichier provisoire avec le troisième fichier "Transit" (1972-73): de là, création d'un second fichier provisoire (P_2);



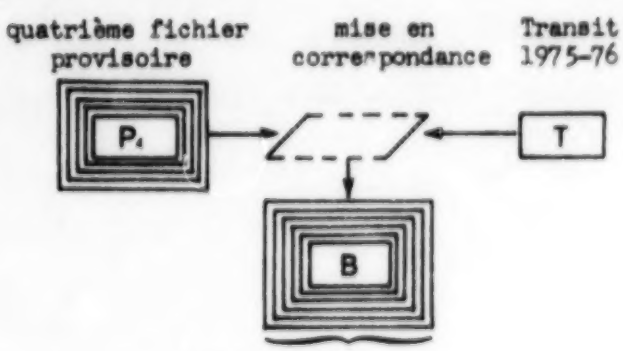
3- mise en correspondance du second fichier provisoire avec le quatrième fichier "Transit" (1973-74): de là, création d'un troisième fichier provisoire (P_3);



4- mise en correspondance du troisième fichier provisoire avec le cinquième fichier "Transit" (1974-75): de là, création d'un quatrième fichier provisoire (P_4);



5- enfin, mise en correspondance du quatrième fichier provisoire avec le sixième fichier "Transit" (1975-76): de là, création d'un dernier fichier qui constitue la banque "Mobilité du Personnel"



Banque "Mobilité du Personnel"

A chaque étape, deux situations sont possibles:

a) l'individu se retrouve inscrit dans les deux fichiers mis en correspondance;

b) l'individu ne se retrouve inscrit que dans un seul des fichiers mis en correspondance.

Si c'est la situation a) qui prévaut, il y a alors jumelage, du type, ici, "inter-Transit". L'information contenue dans le "Transit" (T) est ajoutée au dossier de l'individu et ce dossier est placé dans le nouveau fichier provisoire.

Par contre, si c'est la situation b) qui est en jeu, il n'y a pas de jumelage (inter-Transit). Les dossiers de ces individus sont alors directement incorporés au nouveau fichier.

Il s'ensuit donc que, à la fin de chaque étape, le nombre de dossiers contenus dans le nouveau fichier créé est égal au total des dossiers mis en correspondance (ceux du fichier provisoire[1] + ceux du fichier "Transit"), moins le nombre de ceux qui ont fait l'objet d'un jumelage "inter-Transit".

---

(1) A la première étape, il s'agit du fichier "Transit" 1970-71.

CHAPITRE II

MONTAGE AUTOMATIQUE DE LA BANQUE

Dans le chapitre précédent, nous avons vu la logique des opérations, c'est-à-dire la théorie. Maintenant, nous allons voir comment, en pratique, les choses se sont déroulées, c'est-à-dire comment s'est fait le montage automatique de la Banque "Mobilité du Personnel".

A vrai dire, dès qu'on passe de la théorie à la pratique, on constate que tous les problèmes naissent de l'obligation de choisir des facteurs ou des éléments de montage, pour ainsi dire, qui permettent précisément de réaliser les opérations de triage et de mise en correspondance des fichiers "Transit". Car, suivant les éléments retenus d'une part s'impose un enchaînement des opérations qui n'est pas nécessairement celui que la théorie aurait pu imaginer et d'autre part se détermine le degré de validité de l'opération même.

Laissons pour l'instant la question de la validité du montage automatique de la banque (nous y reviendrons au chapitre III) et examinons brièvement les détails de l'opération de montage.

1. Les éléments de montage

Pour mener à bien l'opération, nous avons utilisé quatre (4) éléments de montage. Ce sont:

1) le No. d'Assurance Sociale - Sexe (N.A.S.S.)
2) le Code Permanent (C.P.): consitué des trois premières lettres du nom, de la première lettre du prénom et de la date de naissance (jour, mois,[1] deux derniers chiffres de l'année).
3) le No. Matricule (N.M.)
4) l'ensemble: "Nom, Prénom, Date de naissance" (N.P.D.)

De ces quatre éléments, c'est le N.A.S.S. qui s'est révélé le plus pratique pour suivre les individus à travers les années. Deux études ont en effet montré la supériorité de cet élément notamment sur le code permanent, les deux autres éléments (N.M. et N.P.D.) étant plutôt des moyens de dernier recours. Dans la première étude, nous avons d'abord cherché à savoir combien de dossiers pouvaient être suivis entre deux années par le N.A.S.S. et le C.P. Les résultats se présentèrent comme suit:

	1970-71/71-72	1971-72/72-73	1972-73/73-74	1973-74/74-75
N.A.S.S.	80 102	83 384	86 698	86 495
C.P.		81 245	85 781	85 522

(1) Dans le cas des personnes du sexe féminin, 50 est ajouté au chiffre du mois.

Puis nous avons essayé de voir combien de dossiers étaient communs aux deux codes, c'est-à-dire pouvaient être suivis par les deux codes, ou étaient uniques à l'un ou à l'autre des codes, c'est-à-dire pouvaient être suivis par l'un et non par l'autre. Les résultats furent les suivants:

	1970-71/71-72	1971-72/72-73	1972-73/73-74	1973-74/74-75
Communs	—	81 061	85 604	85 332
Uniques au N.A.S.S.	—	2 323	1 094	1 163
Uniques au C.P.	—	184	177	190

Il appraît donc que le Numéro d'Assurance Sociale - Sexe présente plus d'intérêt au niveau de la fiabilité.

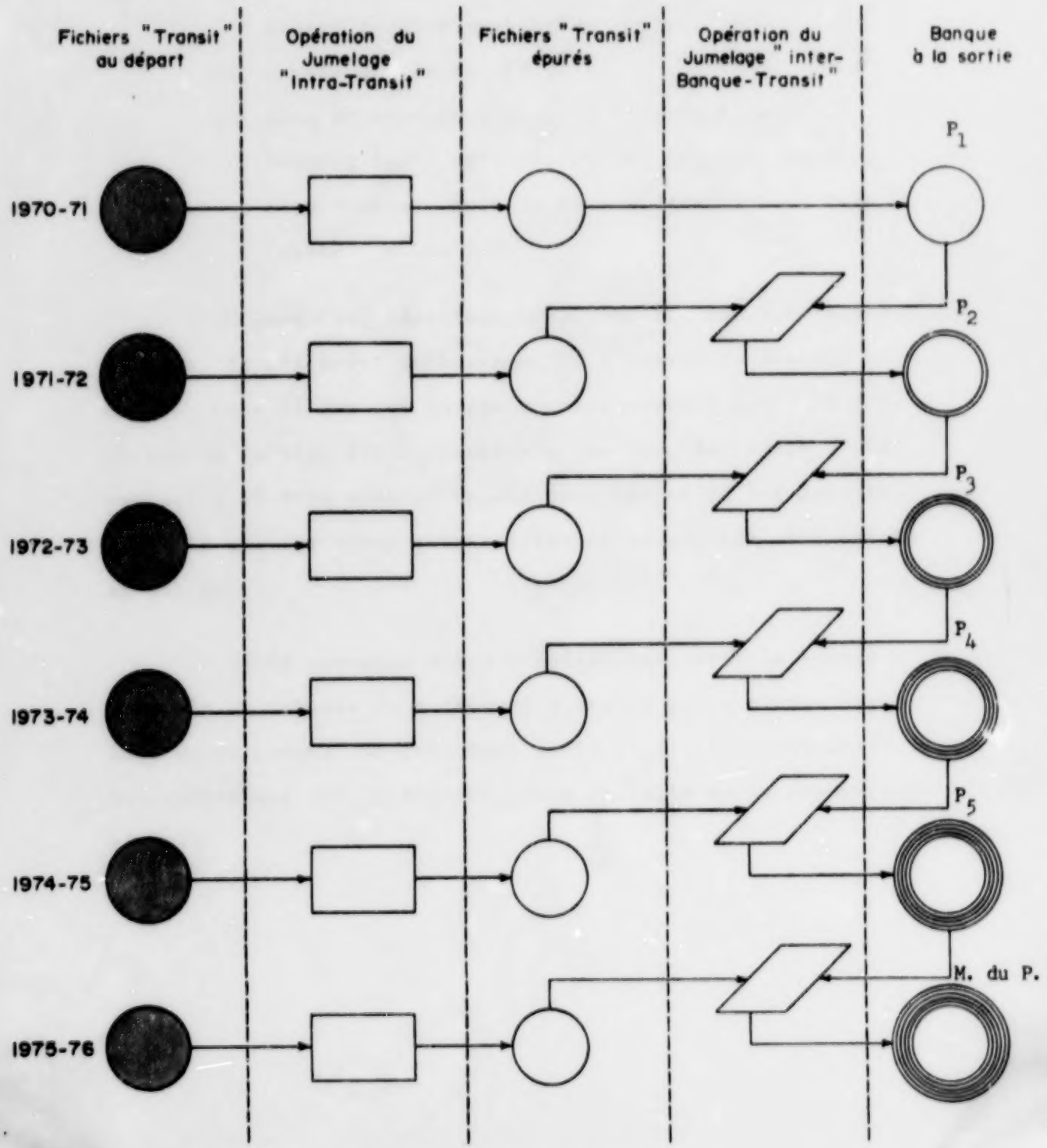
C'est donc au moyen du N.A.S.S. que s'est réalisé le passage des six fichiers "Transit" à la banque "Mobilité du Personnel". C'est-à-dire que c'est le N.A.S.S. qui a servi à la fois à trier les fichiers "Transit" par individu (jumelage intra-Transit) et à mettre en correspondance ces mêmes fichiers (jumelage inter-Transit). Quant aux trois autres éléments de montage (C.P. N.N. et N.P.D.), ils ont été employés uniquement pour tenter de résoudre (jumeler) les dossiers qui avaient présenté un N.A.S.S. défectueux (absent ou incomplet).

2. L'Enchaînement des opérations

Dans le cours du montage automatique de la banque, les deux phases d'opération décrites à la section 1 ont été combinées. L'enchaînement des opérations, schématisé au tableau 1, se présente de la façon suivante:

1) triage du Transit 1970-71; ce dernier devenant automatiquement le premier fichier provisoire (P_1);
2) triage du Transit 1971-72;
3) mise en correspondance du fichier P_1 avec le Transit épuré 1971-72: de là, création du second fichier provisoire (P_2);
4) triage du Transit 1972-73;
5) mise en correspondance du fichier P_2 avec le Transit épuré 1972-73: de là, création du troisième fichier provisoire (P_3);
6) triage du Transit 1973-74;
7) mise en correspondance du fichier P_3 avec le Transit épuré 1973-74: de là, création du quatrième fichier provisoire (P_4);
8) triage du Transit 1974-75;

Tableau 1: Schéma des opérations de montage de la Banque "Mobilité du Personnel", à partir des fichiers "Transit" annuels. Période 1970-71 à 1975-76. Montage sur le code No d'Assurance Sociale-Sexe (N.A.S.S.), ETAPE I



9) mise en correspondance du fichier P_4 avec le Transit épuré 1974-75: de là, création du cinquième fichier provisoire (P_5);

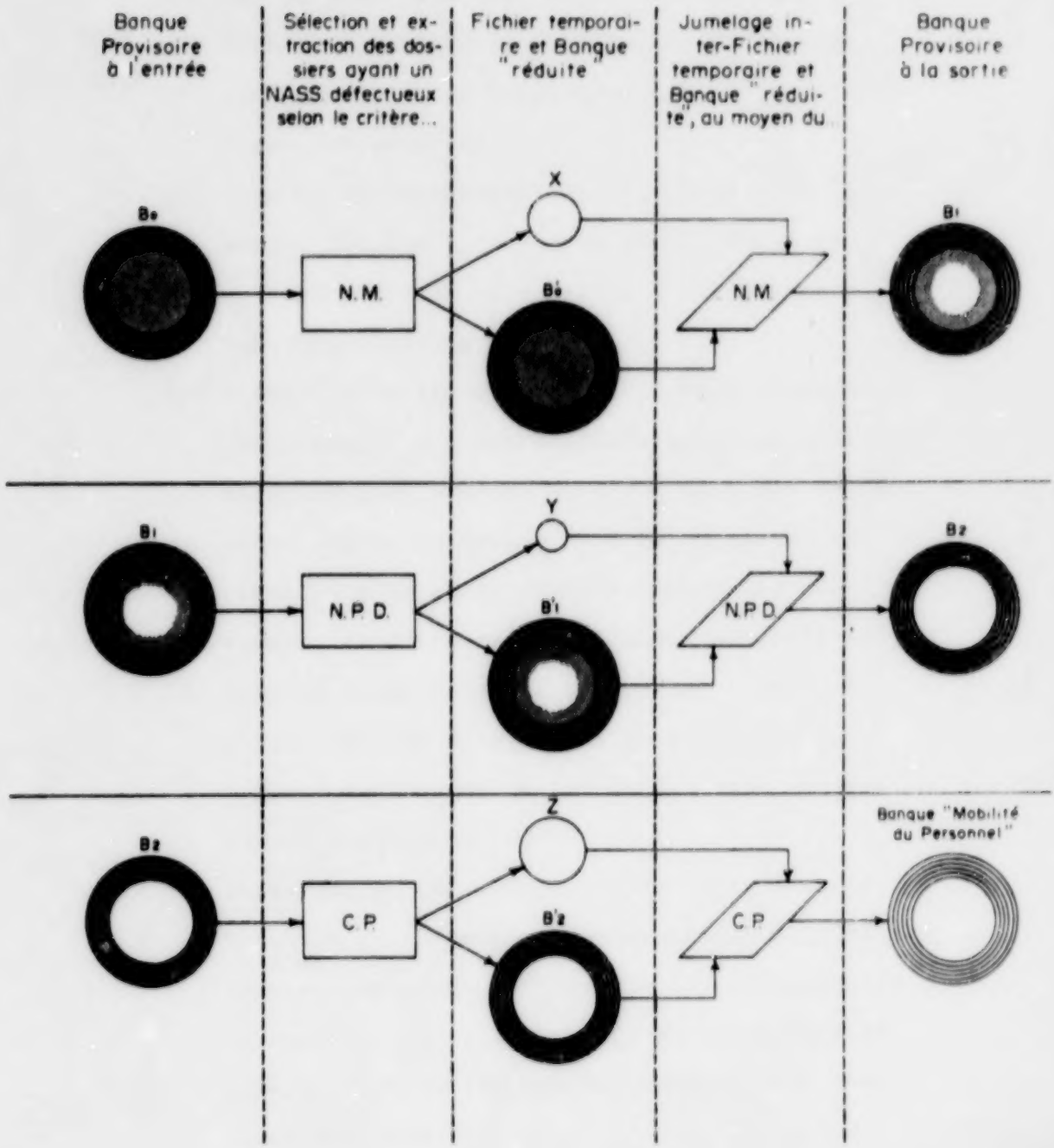
10) triage du transit 1975-76

11) mise en correspondance du fichier P_5 avec le Transit épuré 1975-76: de là, création du dernier fichier, soit la banque "Mobilité du Personnel" (M. du P.).

Toutes ces opérations, répétons-le, ont été faites au moyen de l'élément de montage "No d'Assurance Sociale - Sexe". Dans le cas des dossiers qui n'avaient pas de N.A.S.S. ou que ce dernier était incomplet, une deuxième série d'opérations a dû être mise en branle pour tenter de les jumeler. C'est là que les trois autres éléments de montage sont entrés en action.

Cette nouvelle série d'opérations, dont le schéma apparaît au tableau 2, a démarré à partir de la Banque "Mobilité du Personnel" nouvellement créée (B_0). L'enchaînement des opérations est le suivant (dans la suite de la première série):

TABLEAU 2: Schéma des opérations de jumelage des dossiers ayant un N.A.S.S. défectueux (absent ou incomplet) sur la Banque "Mobilité du Personnel", ETAPE II



12) sélection et extraction, de la Banque B_0, des dossiers ayant un N.A.S.S. défectueux, n'ayant pas de C.P., mais ayant un N.M.: de là, réduction de la Banque B_0 à la Banque B'_0 et création d'un fichier temporaire contenant les dossiers extraits (X);

13) mise en correspondance du fichier X avec la Banque réduite B'_0 au moyen de l'élément de montage N.M.: de là création d'une première Banque provisoire (B_1);

14) sélection et extraction, maintenant de la nouvelle Banque B_1, des dossiers ayant un N.A.S.S. défectueux et n'ayant ni C.P., ni N.M.: de là, réduction de la Banque B_1 à la Banque B'_1 et création d'un autre fichier temporaire contenant les dossiers nouvellement extraits (Y);

15) mise en correspondance du fichier Y avec la Banque réduite B'_1 au moyen de l'élément de montage N.P.D.: de là, création d'une seconde Banque provisoire (B_2);

16) sélection et extraction, cette fois-ci de la Banque B_2, des dossiers ayant un N.A.S.S. défectueux et ayant un C.P.: de là, réduction de la Banque B_2 à la Banque B'_2, et création d'un dernier fichier temporaire contenant les dossiers extraits (Z);

17) enfin, mise en correspondance du fichier Z avec la Banque réduite B'_2 au moyen de l'élément de montage C.P.: de là, création de la Banque finale "Mobilité du Personnel".

3. Résultats du montage

Afin de donner une idée de l'ampleur des opérations, nous avons repris, aux tableaux 3 et 4, le schéma des opérations des tableaux 1 et 2 en y inscrivant le nombre de dossiers impliqués lors des diverses étapes.

C'est ainsi qu'on peut voir que, pour la période 1970-71 à 1975-76, les fichiers "Transit" au départ (non épurés) contenaient un total de 604 173 enregistrements. Les opérations de triage, ou de jumelage intra-Transit, ont permis d'éliminer 5 463 pseudo-individus, réduisant ainsi le nombre d'individus inscrits dans les fichiers "Transit" à 598 710.

Par ailleurs, les cinq opérations de mise en correspondance des "Transit" sur le N.A.S.S. ont permis 430 551 jumelages (inter-Transit), établissant par conséquent que les 598 710 enregistrements appartenaient en réalité à 168 159 individus.

Tableau 3: Schéma des opérations de montage de la Banque "Mobilité du Personnel", à partir des fichiers "Transit" annuels. Période 1970-71 à 1975-76. Montage sur le code No d'Assurance Sociale-Sexe (N.A.S.S.), ETAPE I Nombre de dossiers impliqués

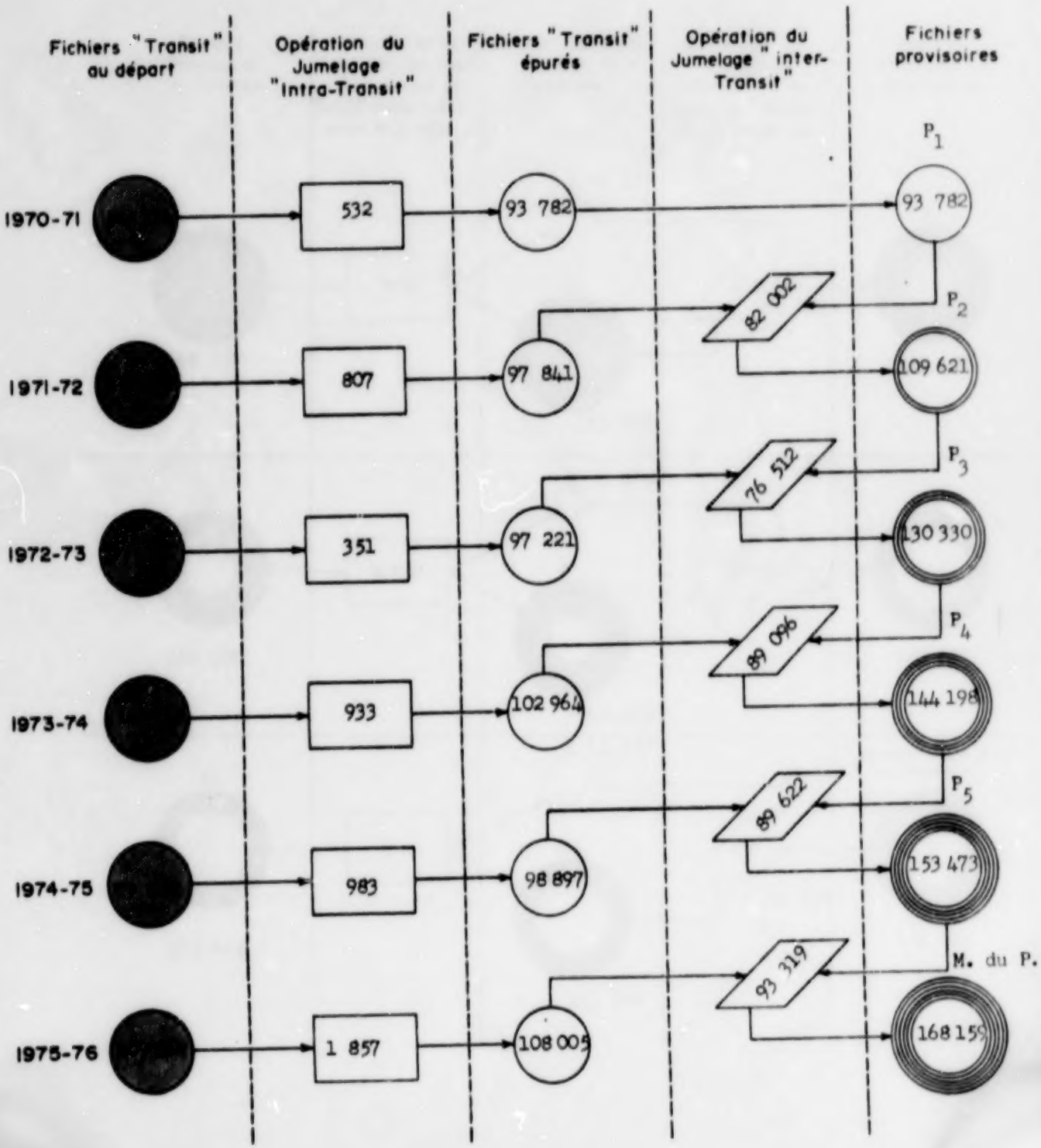
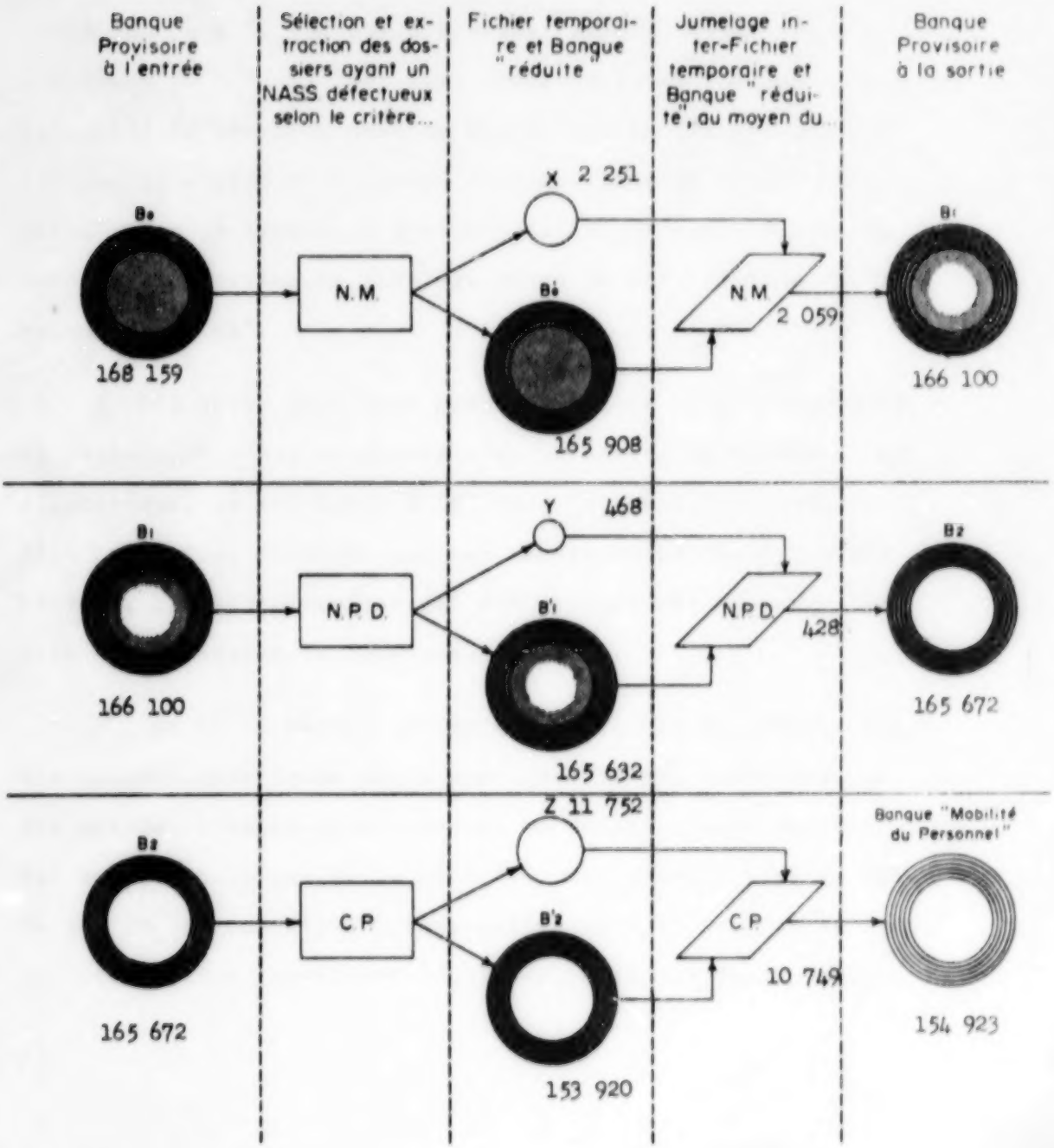


TABLEAU 4: Schéma des opérations de jumelage des dossiers ayant un N.A.S.S. défectueux (absent ou incomplet) sur la Banque "Mobilité du Personnel", ETAPE II. Nombre de dossiers impliqués



Finalement, la seconde série d'opérations concernant les enregistrements qui n'avaient pas un N.A.S. valide a permis 13 256 jumelages supplémentaires: 2 059 sur l'élément "N.M.", 428 sur l'élément "N.P.D." et 10 749 sur l'élément "C.P." Il faut bien comprendre, au tableau 4, que le nombre de dossiers dans la banque provisoire est constitué par le nombre de dossiers contenus dans la banque "réduite" plus le nombre de dossiers qui n'ont pas été jumelés lors de l'opération du jumelage entre le Fichier temporaire et la Banque "réduite".

De sorte que, tout compte fait, la banque "Mobilité du Personnel" s'est retrouvée avec un nombre de dossiers, ou d'individus, se chiffrant à 154 923. C'est-à-dire que, durant la période couverte par les années scolaires 1970-71 à 1975-76, 154 923 individus ont été enregistrés au moins une fois comme membres du personnel scolaire.

Si on se réfère au tableau 5, on est en mesure de prendre connaissance de la façon dont ces 154 923 individus ont été actifs, c'est-à-dire inscrits au fichier, pour chacune des années scolaires de la période. On remarque tout de suite que 57 322 individus furent actifs tout au long de la période. Ce chiffre représente 37,0% des individus de la banque.

TABLEAU 5: Etat de l'activité des individus de la Banque "Mobilité du Personnel" pour les six années scolaires de la période 1970-71 à 1975-76 (activité étant synonyme de "inscription au fichier")

Années scolaires						Nombre d'individus	%
1970-71	1971-72	1972-73	1973-74	1974-75	1975-76		
					X	14 387	9,29
				X		2 705	1,75
				X	X	6 210	4,01
			X			4 189	2,70
			X		X	880	0,57
			X	X		1 634	1,05
			X	X	X	5 887	3,80
		X				2 281	1,47
		X			X	205	0,13
		X		X		91	0,06
		X		X	X	264	0,17
		X	X			1 237	0,80
		X	X		X	394	0,25
		X	X	X		716	0,46
		X	X	X	X	5 569	3,59
	X					2 684	1,73
	X				X	190	0,12
	X			X		39	0,03
	X			X	X	231	0,15
	X		X			155	0,10
	X		X		X	50	0,03
	X		X	X		81	0,05
	X		X	X	X	429	0,28
	X	X				1 107	0,72
	X	X			X	148	0,10
	X	X		X		61	0,04
	X	X		X	X	200	0,13
	X	X	X			1 023	0,66
	X	X	X		X	361	0,23
	X	X	X	X		746	0,48
	X	X	X	X	X	7 003	4,52
X(1)						7 132	4,60
X					X	332	0,21
X				X		80	0,05
X				X	X	368	0,24
X			X			207	0,13
X			X		X	62	0,04
X			X	X		96	0,06
X			X	X	X	445	0,29
X		X				229	0,15
X		X			X	44	0,03
X		X		X		11	0,01
X		X		X	X	56	0,04
X		X	X			184	0,12
X		X	X		X	63	0,04

(1) le "X" indique que les individus étaient actifs (inscrits au fichier) pour l'année en question.

TABLEAU 5: Etat de l'activité des individus de la Banque "Mobilité du Personnel" pour les six années scolaires de la période 1970-71 à 1975-76 (activité étant synonyme de "inscription au fichier")

Années scolaires						Nombre d'individus	%
1970-71	1971-72	1972-73	1973-74	1974-75	1975-76		
					X	14 387	9,29
				X		2 705	1,75
				X	X	6 210	4,01
			X			4 189	2,70
			X		X	880	0,57
			X	X		1 634	1,05
			X	X	X	5 887	3,80
		X				2 281	1,47
		X			X	205	0,13
		X		X		91	0,06
		X		X	X	264	0,17
		X	X			1 237	0,80
		X	X		X	394	0,25
		X	X	X		716	0,46
		X	X	X	X	5 569	3,59
	X					2 684	1,73
	X				X	190	0,12
	X			X		39	0,03
	X			X	X	231	0,15
	X		X			155	0,10
	X		X		X	50	0,03
	X		X	X		81	0,05
	X		X	X	X	429	0,28
	X	X				1 107	0,72
	X	X			X	148	0,10
	X	X		X		61	0,04
	X	X		X	X	200	0,13
	X	X	X			1 023	0,66
	X	X	X		X	361	0,23
	X	X	X	X		746	0,48
	X	X	X	X	X	7 003	4,52
X(1)						7 132	4,60
X					X	332	0,21
X				X		80	0,05
X				X	X	368	0,24
X			X			207	0,13
X			X		X	62	0,04
X			X	X		96	0,06
X			X	X	X	445	0,29
X		X				229	0,15
X		X			X	44	0,03
X		X		X		11	0,01
X		X		X	X	56	0,04
X		X	X			184	0,12
X		X	X		X	63	0,04

(1) le "X" indique que les individus étaient actifs (inscrits au fichier) pour l'année en question.

TABLEAU 5 (fin): **Etat de l'activité des individus de la Banque "Mobilité du Personnel" pour les six années scolaires de la période 1970-71 à 1975-76 (activité étant synonyme de"inscription au fichier")**

Années scolaires						Nombre d'individus	%
1970-71	1971-72	1972-73	1973-74	1974-75	1975-76		
X		X	X	X		152	0,10
X		X	X	X	X	1 160	0,75
X						5 773	3,73
X	X				X	445	0,29
X	X			X		110	0,07
X	X			X	X	749	0,48
X	X		X			391	0,26
X	X		X		X	161	0,10
X	X		X	X		200	0,13
X	X		X	X	X	1 438	0,93
X	X	X				4 219	2,72
X	X	X			X	599	0,39
X	X	X		X		178	0,11
X	X	X		X	X	852	0,55
X	X	X	X			5 474	3,53
X	X	X	X		X	1 452	0,94
X	X	X	X	X		3 782	2,44
X	X	X	X	X	X	57 322	37.00
						154 923	100,00

Egalement, on note que 33 378 individus (21,5%) ne furent actifs qu'une seule année. Enfin, on peut calculer que 55,8% des individus peuvent être suivis à compter du début de la période (excluant la proportion de ceux qui ne sont actifs qu'en 1970-71 (4,60%). Ce chiffre constitue un pourcentage fort intéressant.

CHAPITRE III

VALIDITE DU PROCESSUS DE MONTAGE

La première question ou, devrions-nous dire, le premier doute qui vient à l'esprit se pose certes au niveau de la validité d'un tel processus. Dans quelle mesure, en effet, le fait d'avoir deux enregistrements possédant les mêmes "Numéros d'Assurance Sociale - Sexe", pour ne nommer que cet élément, garantit que ces enregistrements appartiennent bien à la même personne? C'est là, hors de tout doute, une question cruciale.

Pour contrôler la validité du processus de jumelage, nous avons créé deux variables spéciales:

1) le code de résolution
2) le code d'authenticité

1. Le code de résolution

Le code de résolution nous permet de savoir exactement de quelle façon chacun des 154 923 dossiers de la banque a été traité au cours de l'opération de montage automatique de la banque.

C'est ainsi que, grâce à cette variable, nous sommes en mesure d'affirmer que 71,2% des dossiers de la banque ont

été jumelés au moyen du seul élément "N.A.S.S.", que 0,36% l'ont été au moyen du seul élément "C.P.", 0,19% au moyen du seul élément "N.M." et, enfin, 0,04% au moyen du seul élément "N.P.D." (voir tableau 6).

En outre, on note que 21,21% des dossiers n'ont fait l'objet d'aucun jumelage. C'est donc dire que nous avons 32 855 individus qui n'ont été actifs qu'une seule année et qui, en plus, ne possédaient alors qu'un seul enregistrement à leur nom (sinon ils auraient été impliqués dans une opération de jumelage "intra-Transit").

TABLEAU 6: Statistiques sur l'utilisation des éléments de montage concernant les dossiers de la Banque "Mobilité du Personnel", au cours des opérations de montage automatique de la Banque

ELEMENTS DE MONTAGE				Nombre de dossiers	%
N.A.S.S.(1)	C.P.	N.M.	N.P.D.		
			X(2)	32 855	21,21
				69	0,04
		X		286	0,19
		X	X	18	0,01
	X			559	0,36
	X		X	7	0,00
	X	X		21	0,02
	X	X	X	2	0,00
X				110 304	71,20
X			X	239	0,15
X		X		912	0,59
X		X	X	70	0,05
X	X			9 456	6,10
X	X		X	17	0,01
X	X	X		104	0,07
X	X	X	X	4	0,00
				154 923	100,00

(1) N.A.S.S.: "Numéro d'Assurance Sociale - Sexe"
C.P.: Code Permanent
N.M.: Numéro Matricule
N.P.D.: "Nom - Prénom - Date de naissance"

(2) Le "X" indique que l'élément en question a été utilisé.

On remarque aussi que 6,10% des dossiers ont requis les deux éléments "N.A.S.S." et "C.P.". Pour la presque totalité de ces dossiers, le C.P. a servi à faire le lien au niveau de l'année 1972-73. C'est en cette année, en effet, qu'a été créé le code permanent. Le S.I.M.E.Q. lui a alors accordé, un peu aux dépens du N.A.S.S., la priorité au niveau des tests de cohérence et de validité. Il en est donc résulté que plusieurs dossiers se sont retrouvés avec un N.A.S.S. défectueux. Pour les années suivantes, la situation revint cependant à la normale et le N.A.S.S. reprit son "autorité".

Enfin, on peut calculer que 4 dossiers ont requis les quatre éléments de montage, que 193 dossiers en ont nécessité trois (0,12%) et que 10 653 dossiers en ont nécessité deux (6,88%).

2. <u>Le code d'authenticité</u>

Ce code nous permet d'évaluer la fiabilité du processus de jumelage, et ce pour chaque individu (dossier) de la banque. Trois valeurs sont possibles:

- "blanc" (ϐ): jumelage inexistant
- 1 : jumelage fiable
- 2 : Jumelage douteux

La première valeur (ϐ) signale que le dossier en question n'a jamais fait l'objet d'un jumelage quelconque (ni "intra, ni

"inter" Transit). Ce sont ici nos 32 885 dossiers de la première ligne du tableau 6.

La valeur "1" signifie que, à chaque étape du montage au moyen du N.A.S.S., les dossiers jumelés avaient, dans tous les cas, des C.P. identiques. Nous avons donc émis l'hypothèse que, si deux dossiers avaient les mêmes N.A.S.S. et les mêmes C.P., la probalité d'appartenir à la même personne était égale à 1.

La valeur "2", elle, fait état de deux situations possibles:

situation 1: à un moment donné, dans le cours du montage au moyen du N.A.S.S., le dossier en question a été jumelé avec un autre dossier, alors que leurs C.P. n'étaient pas identiques:

situation 2: le dossier a été jumelé au moyen d'un élément de montage autre que le N.A.S.S., soit au moyen du C.P., soit au moyen du N.M. ou du N.D.P.

La distribution des dossiers entre ces différentes valeurs possibles se présente comme suit:

Valeur possible	Nombre de dossiers	%
∅	32 855	21,21
1	102 657	66,26
2 situation (1)	6 175	3,99
situation (2)	13 236	8,54
	154 923	100,00

En ce qui concerne les jumelages douteux, il faut bien voir que ce ne sont pas nécessairement de mauvais jumelages. Par exemple, pour ceux faits au moyen du N.A.S.S., il s'agissait qu'il y eût une inversion entre le jour et le mois de naissance (10/12/47 et 12/10/47) ou une différence dans le prénom (ex.: Berthe et Sr. Berthe) pour que les C.P. fussent différents et que le code d'authenticité devînt "2". Quant aux jumelages faits au moyen du C.P., au moyen du N.M. ou du N.P.D., ils étaient automatiquement classés "2". Aussi nous est-il apparu plus juste d'examiner un peu plus à fond les dossiers de code d'authenticité "2".

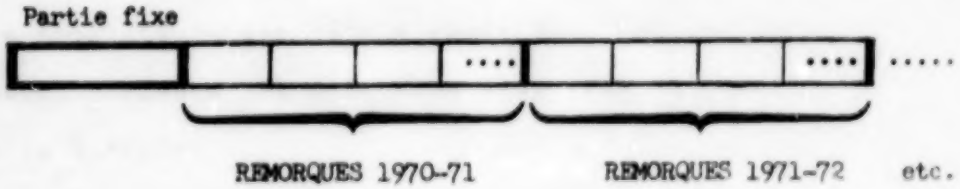
Pour ce faire, nous avons fait écrire, pour chaque jumelage douteux, le N.A.S.S., le C.P., le nom, le prénon, le N.M. et la date de naissance des deux individus jumelés. Et même, dans certains cas, nous avons fait sortir le dossier complet des individus jumelés. C'est ainsi que nous avons

pu établir que 98% des dossiers de code d'authenticité "2" devraient, tout compte fait, avoir la valeur "1"; en fait 2% seulement des 19 411 dossiers de code d'authenticité "2" ont fait l'objet d'un doute impossible à dissiper. Si on calcule que 2% de cas douteux représentent 0,25% du total des dossiers de la banque, il faut donc se réjouir du résultats et reconnaître que nous disposons là d'une banque grandement fiable.

CONCLUSION

La puissance de notre banque "Mobilité du Personnel" vient donc essentiellement du fait que tout est ramené en fonction de l'individu, et ce, dans le champ d'une vision continue des choses tout au long de la période 1970-71 à 1975-76.

Au lieu de traiter 604 173 dossiers durant les six années de la période, la banque en traite seulement 154 923 qui appartiennent tous à des individus différents. Les individus possèdent donc chacun un seul dossier où toute l'information est consignée. D'ailleurs, la structure même de la banque a été pensée dans ce sens. En voici une illustration.



On a tout d'abord, une partie fixe établissant l'identité de l'individu, puis une série de "remorques" contenant l'information annuelle pour chacune des années.

Il n'est pas besoin de nous étendre longuement sur les multiples avantages qu'une telle banque représente.

Soulignons, entre autres, qu'elle élimine tous les cas de "double" (un individu enregistré deux fois une même année). D'autre part, la structure de la banque permet de contrôler presque parfaitement le phénomène des faux mouvements, puisque les données relatives à une année donnée peuvent être comparées à celles des années antérieures ou postérieures. En outre, c'est une banque qui se manie très bien, vu que toute l'information est standardisée sur les six ans de la période.

Il y a donc un pas énorme entre les fichiers de base "Personnel-Maître", miroirs de situations du moment, et cette banque de données qui éclaire une tranche d'histoire si petite soit-elle. Avec cette possiblité de suivre les membres du personnel scolaire individu par individu, la vision sur la mobilité de ces gens s'ouvre donc sur des horizons nouveaux, horizons que nous nous proposons de sonder dans les mois à venir.

ANNEXE

LISTE DES RAPPORTS D'ETAPE

COUVRANT LE PROJET "ELMOPE" '70-75"

RAPPORTS D'ETAPE

. LA HAYE, Jacques — "ELMOPE '70-75". Etude longitudinale de la mobilité des personnels de l'enseignement sur la période 1970-75. Rapport d'étape no 1, Démographie scolaire, Document de travail D.T. 9.36, Direction générale de la planification,

. LA HAYE, Jacques — "ELMOPE '70-75", Montage du fichier "Mobilité du personnel": choix des variables, standardisation des variables, traitement informatique des variables, Rapport d'étape no 2, Démographie scolaire, Document de travail D.T. 9.38, Direction générale de la planification, M.E.Q., octobre 1976, 76 p.

. LA HAYE, Jacques — "ELMOPE '70-75", Montage du fichier "Mobilité du personnel", analyse des variables: service et code de service, fonction et code de fonction, Rapport d'étape no 3, Démographie scolaire, Document de travail D.T. 9.39, Direction générale de la planification, M.E.Q., octobre 1976, 41 p.

. FERLAND, Robert — "ELMOPE '70-75", Description de six fichiers annuels "TRANSIT" 1970-71 à 1975-76 et du fichier "MOBILITE DU PERSONNEL" 1970-75, Rapport d'étape no 4, Démographie scolaire, Document de travail D.T. 9.41, Direction générale de la planification, M.E.Q. octobre 1976, 179 p.

L'ÉDITEUR OFFICIEL DU QUÉBEC
SERVICE DE LA REPROGRAPHIE
Septembre 1977